Computational Numerical Statistics
Assignment 11

Migla Miskinyte Reis

26.12.20

Exercise 4. (0.1 points) **Points:** _______

Let X_1,...,$X_n \sim Pois(\theta)$. Show that $Gamma$(a, b) is a natural conjugate prior for θ.

Assuming a parameter $\alpha \in (0, \infty)$ from $Pois$ and prior distribution for α is $Gamma$(a, b) distribution, we can write a prior for α :

$$f_\alpha(\theta) = \frac{b^a}{\Gamma(a)}\theta^{(a-1)}e^{-b\theta}$$

and a likelihood function:

$$f_{X|\alpha}(x_1, ..., x_n|\alpha) = \prod_{i=1}^{n}\frac{\theta^{x_i}e^{-\theta}}{x_i!}$$

If we ignore constants that do not depend on θ, the posterior distribution of α parameter given that:

$$X_1{=}x_1,...,X_n{=}x_n$$

is

$$f_{\alpha|X}(\alpha|x_1, ..., x_n) = f_{X|\alpha}(x_1, ..., x_n|\alpha)f_\alpha(\theta) =$$

$$= \prod_{i=1}^{n}\theta^{x_i}e^{-\theta}\theta^{(a-1)}e^{-b\theta}$$

$$= \theta^{c+a-1}e^{-(n+b)\theta},$$

where c is equal to x_1+...+x_n. We can see that this equation is the same as probability density function of the gamma distribution $Gamma$(c+a,n+b), so that means that the posterior distribution of parameter α must be $Gamma$(c+a,n+b). Because the prior and posterior are both gamma distributions, the gamma distribution is a natural conjugate prior in the Poisson model.

Exercise 6. (0.1 points) **Points:** _______

Consider a drug to be given for the relief of chronic pain and that experience with similar compounds has suggested that response rates, say θ, between 0.2 and 0.6 could be feasible. Let Y be the number of patients that experienced pain relief (positive response) in a sample of n treated patients. One has $Y \sim Bin(n, \theta)$.

According to the prior information above, a prior for θ should be chosen such that it has mean $\mu = \frac{0.2+0.6}{2}$=0.4 and such that the standard deviation σ is so that $\mu \pm 2$ x σ gives us the interval [0.2, 0.6], i.e, $\sigma = 0.1$. Assume that n = 20 volunteers were treated with the compound and that out of those we observed y = 15 positive responses to treatment.

6.1 Given that the Beta distribution is a conjugate prior for the Binomial model, identify the prior distribution's hyperparameters.

Given $\theta \sim Beta$(a,b) is our prior for Binomial model, we know that prior mean is:

$$\mathrm{E}(\theta) = \frac{a}{a+b}=0.4$$

From here:

$$b=1.5a$$

We also know that standard deviation $\sigma = 0.1$ and variance for prior Beta distribution is:

$$\mathrm{Var}(\theta) = \frac{ab}{(a+b)^2(a+b+1)}=SD^2=0.01$$

We substitute b=1.5a and solve the previous equation, where we find a=$\frac{46}{5}$=9.2, so b = 13.8. From here we can see that our prior Beta(a,b) is approximately Beta (9,14).

6.2

Derive the posterior distribution p(θ|y). We know that in order to calculate posterior distribution for the Beta-Binomial case, we can derive a very general relationship between the likelihood, prior, and posterior: $p(\theta|X) \propto p(\theta)p(X|\theta)$. Given the Binomial likelihood up to proportionality (ignoring the constant):$\theta^y(1-\theta)^{n-y}$ and given the prior, also up to proportionality,$\theta^{a-1}(1-\theta)^{b-1}$, we can write:

$$\theta^y(1-\theta)^{n-y}\theta^{a-1}(1-\theta)^{b-1} = \theta^{a+y-1}(1-\theta)^{b+n-y-1}$$

From this given Binomial likelihood $(n, y|\theta)$ and prior Beta(a,b), our posterior will be:

$$Beta(a+y, b+n-y)$$

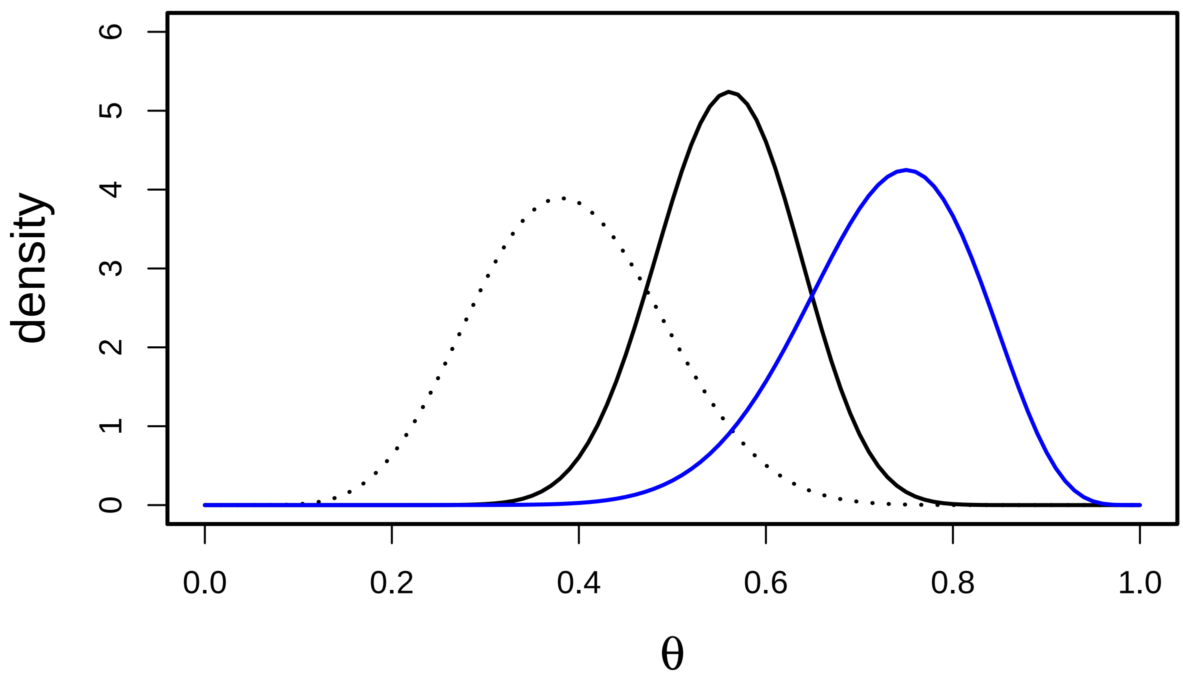
So, in is simply $Beta(9+15, 14+20-15) = Beta(24, 19)$.

6.3 Given the observed data, plot the likelihood versus the prior and the posterior as in the class **Example 2**.

```
n = 20; y = 15; a = 9; b = 14

# priori Beta(9,14)
curve(dbeta(x,a,b),from=0,to=1,lwd=2,
type="l",ylab="density",
xlab=expression(theta),
cex.lab=1.5,cex.main=2,
lty=3,ylim=c(0,6))
# posteriori Beta(24,19)
curve(dbeta(x,a+y,b+n-y),add=T,lwd=2,
type="l",ylab="density",cex.lab=1.5,
```

```r
cex.main=2)
# proportional likelihood Beta(16,6)
curve(dbeta(x,y+1,n-y+1),add=T,col=4,lwd=2)
legend(0.25, 14,
legend=c("prior","posterior","likelihood"),
col=c(1,1,4),lty=c(3,1,1),cex=1,
lwd=c(1,2,2))
box(lwd=2)
```



6.4

Compute the Bayes estimate of θ for the quadratic loss and compare it with the ML estimate.

```r
theta.mle = y/n; theta.mle; #calculating ML estimate
```

```
## [1] 0.75
```

```r
#Prior mean
e.theta.prior = a/(a+b); e.theta.prior #calculating theta prior
```

```
## [1] 0.3913043
```

```r
e.theta.post = (a+y)/(a+b+n); e.theta.post #theta postoreior
```

```
## [1] 0.5581395
```

Now we will compute Bayes estimate of θ for the quadratic loss. Our prior for $\theta \sim Beta(9{,}14)$ and posterior $\theta|y \sim Beta(24{,}19)$.

$$\hat{\theta}_B = \mathrm{E}(\theta|\mathrm{y}) = \frac{24}{24+19} = 0.058$$

Or, calculating in R:

```r
a=24; b=19
# Bayes estimator with quadratic loss
e.theta = a/(a+b); e.theta
```

```
## [1] 0.5581395
```

Bayes estimate of θ for the quadratic loss is 0.558, while ML estimate is 0.75.

6.5 Compute both the 90% CC and 90%-HPD intervals for θ. We could also display the 90% credible interval, the range over which we are 90% certain the true value of θ lies, given the data and model.

```r
qbeta(c(0.05,0.95),shape1=24,shape2=19)
```

```
## [1] 0.4332790 0.6798878
```

```r
library(HDInterval)
HDP = hdi(qbeta, 0.90, shape1=24, shape2=19); HDP
```

```
##     lower     upper
## 0.4351697 0.6817033
## attr(,"credMass")
## [1] 0.9
```

6.6

Test the researcher's hypothesis that the percentage of treated individuals that experience a positive result is less or equal than 35%.

We want to test the following hypothesis:

$$H_0\text{:Y<=0.35 vs } H_1\text{:Y>0.35.}$$

For this we need to calculate the prior odds:

```r
prior.odds = pbeta(0.35,shape1=9, shape2=14)/(1-pbeta(0.35,shape1=9, shape2=14)); prior.odds
```

```
## [1] 0.5465314
```

And then compute posterior odds:

```r
posterior.odds = pbeta(0.35,shape1=24, shape2=19)/(1-pbeta(0.35,shape1=24, shape2=19)); posterior.odds
```

```
## [1] 0.002752952
```

We then calculate the Bayes factor for H_0 hypothesis:

```r
BF = posterior.odds/prior.odds; BF
```

```
## [1] 0.005037134
```

We reject our null hypothesis assuming that we test at 95% significance level.

6.7

What's the expected number of treated patients that experience pain relief in a future sample of m = 5?

Given the posterior Beta(9,14) distribution as our prior for now, one now wishes to predict the number of treated patients that experience pain relief with m=5, for this we need to compute E(z|y). We will use the R library LearnBayes to compute these probabilities at once via function pbetap().

```r
library(LearnBayes)
ab = c(24,19) # posterior (c,d) values are now the prior (a,b) values for pbetap()
m = 5 # new sample size
z = 0:m # values of z for which we want to compute the probabilities
pred = pbetap(ab,m,z) # call the R function
pred
```

```
## [1] 0.02193634 0.11445044 0.26011465 0.32204670 0.21738153 0.06407034
```

```r
e.Z = sum(z*pred); e.Z
```

```
## [1] 2.790698
```

The expected mean of the number treated patients that experience pain relief in a future sample of m = 5 inspected individuals is E(Z) = 2.8.